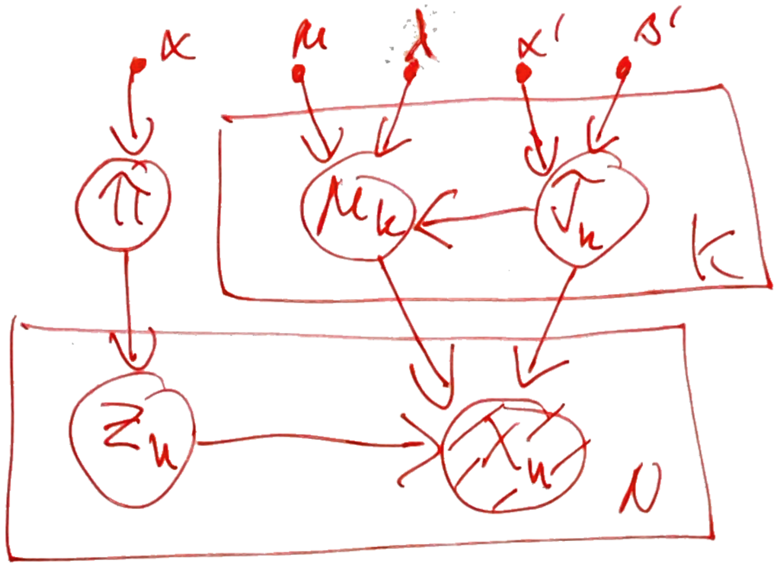
VI GMM

GM



$X = \{x_n : n \in [N]\}$

$Z = \{z_n : n \in [N]\}$

$M = \{\mu_k : k \in [K]\}$

$T = \{\tau_k : k \in [K]\}$

CPDs (w/o hyperparam.)

$$\pi \mid \alpha \sim \mathrm{Dir}(\alpha) = \prod_{k=1}^{K} \pi_k^{\alpha_k - 1}$$

$$\tau_k \mid \alpha', \beta' \sim \mathrm{Ga}(\alpha', \beta') = \frac{\beta'^{\alpha'}}{\Gamma(\alpha')} \tau_k^{\alpha' - 1} e^{-\beta' \tau_k}$$

$$\mu_k \mid \tau_k \sim N(\mu, (\lambda \tau_k)^{-1}) = \sqrt{\frac{\lambda \tau_k}{2\pi}} e^{-\frac{\lambda \tau_k}{2}(\mu_k - \mu)^2}$$

$$z_n \mid \pi \sim \mathrm{Cat}(\pi)$$

$$x_n \mid z_n = k, M, T \sim N(x_n \mid \mu_k, \tau_k) = \sqrt{\frac{\tau_k}{2\pi}} e^{-\frac{\tau_k}{2}(x_n - \mu_k)^2}$$

Joint

$$p(X, Z, \pi, \mu, T) = p(X \mid Z, \mu, T)\, p(Z \mid \pi)\, p(\pi)\, p(\mu \mid T)\, p(T)$$

VI ass.

$$q(Z, \pi, \mu, T) = q(Z)\, q(\pi, \mu, T)$$

Notation

$$z_n \rightarrow z_{n1}, \ldots, z_{nK}$$

where

$$z_n = k \iff z_{nk} = 1, \; z_{nk'} = 0 \quad \forall k' \neq k$$

Complete like.

$$p(X, Z \mid \pi, \mu, T)$$

$$= \prod_n p(x_n \mid z_n, \mu, T)\, p(z_n \mid \pi)$$

$$= \prod_{n,k} \left(p(x_n \mid \mu_k, T_k)\, p(z_{nk} = 1 \mid \pi) \right)^{z_{nk}}$$

$$= \prod_{n,k} \left(\mathcal{N}(x_n \mid \mu_k, T_k^{-1})\, \pi_k \right)^{z_{nk}}$$

Update equations

$\log q^*(Z)$

$$\stackrel{+}{=} \mathbb{E}_{\pi,\mu,\tau}\left[\log p(x, z \mid \pi, \mu, \tau)\right]$$

$$= \sum_{n,k} z_{n,k} \left(\mathbb{E}_{\mu,\tau}\left[\log \mathcal{N}(x_n \mid \mu_k, \tau_k)\right] + \mathbb{E}_{\pi}\left[\log \pi_k\right]\right)$$

$$\stackrel{+}{=} \sum_{n,k} z_{n,k} \underbrace{\left(\frac{\mathbb{E}_{\tau_k}[\log \tau_k]}{2} - \frac{\mathbb{E}_{\tau_k}[\tau_k]}{2}\mathbb{E}_{\mu_k}\left[(x_n-\mu_k)^2\right] + \mathbb{E}_{\pi}\left[\log \pi_k\right]\right)}_{\log \rho_{n,k}}$$

$\boxed{\text{to comp. se below}}$

$$= \sum_{n,k} z_{n,k} \log \rho_{n,k}$$

So $q^*(z_n) = \prod_k r_{n,k}^{z_{n,k}}$

where

$$r_{n,k} = \frac{e^{\rho_{n,k}}}{\sum_k e^{\rho_{n,k}}}$$

$\ln q^*(\pi, M, T)$

$$= E_z[\log p(X, Z | \pi, M, T)] + \log p(\pi) + \log p(M, T)$$

$$= \sum_{n,k} E_z[z_{n,k}] \log N(x_n | \mu_k, T_k^{-1})$$

$$+ \sum_{n,k} E_z[z_{n,k}] \log \pi_k + \sum_k (\alpha_k - 1) \log \pi_k$$

$$+ \sum_k \log p(\mu_k, T_k)$$

No term with π and μ_k or T_k

— " — k and k' $(k \neq k')$

$\Rightarrow$ we can write $q(\pi, M, T) = q(\pi) \prod_k q(\mu_k, T_k)$

Notice also

$$E_z[z_{n,k}] = r_{n,k} \quad \text{(see above)}$$

Terms of (i) with μ_k and τ_k gives

$$\log q^*(\mu_k, \tau_k)$$

$$\stackrel{+}{=} \sum_n r_{n,k} \left(\overset{G}{\frac{\log \tau_k}{2}} - \overset{G}{\frac{\tau_k}{2}} (x_n \overset{N}{-} \overset{N}{\mu_k})^2 \right)$$

$$+ \overset{N}{\frac{\log \tau_k}{2}} - \frac{\lambda \tau_k}{2} (\overset{N}{\mu_k} \overset{N}{-} \overset{G}{\mu})^2$$

$$+ \overset{G}{(\alpha - 1) \log \tau_k} - \overset{G}{\beta' \tau_k}$$

So $q^*(\mu_k | \tau_k) = N(\mu_k | \mu^*, \tau^*)$ where

$$\tau^* = \left(\tau_k \sum_n r_{n,k} \right) + \lambda \tau_k = \tau_k \left(\lambda + \sum_n r_{n,k} \right)$$

$$\mu^* = \frac{\left(\tau_k \sum_n r_{n,k} x_n \right) + \lambda \tau_k \mu}{\tau^*}$$

and

$q^*(\tau_k) = Ga(\lambda_k | \alpha^*, \beta^*)$ where

$$\alpha^* = \alpha + \sum_n r_{n,k}$$

$$\beta^* = \beta' + \frac{\lambda}{2} \mu^2 + \frac{1}{2} \sum_n r_{n,k} x_n^2$$

$\log q^*(\pi)$

$$\stackrel{+}{=} E_z\left[\log p(z|\pi) + \log p(\pi)\right]$$

$$= E_z\left[\sum_{n,k} z_{nk} \log \pi_k + (\alpha_k - 1) \log \pi_k\right]$$

$$= \sum_k \left((\alpha_k - 1) + \sum_n r_{nk}\right) \log \pi_k$$

So $\quad q^*(\pi) = \mathrm{Dir}(\alpha^*)$

where $\quad \alpha^* = \alpha_k + \sum_n r_{n,k}$

For $\log \rho_{nk}$

$E_{\tau_k}[\log \tau_k]$ use digamma

$E_{\mu_k}[(x_n - \mu_k)^2]$ use $V[X] = E[X^2] - E[X]^2$

$E_\pi[\log \pi_k]$ use digamma

$E_{\tau_k}[\tau_k]$ is α^*/β^*